# Tree Crown Extraction using a Three State Markov Random Field

Xavier Descombes, Eugene Pechersky

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

# *Tree Crown Extraction using a Three State Markov Random Field*

Xavier Descombes — Eugène Pechersky

**N° 5982**

Septembre 2006

Thème COG

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5982--FR+ENG

# Tree Crown Extraction using a Three State Markov Random Field

Xavier Descombes , Eugène Pechersky *

Thème COG — Systèmes cognitifs
Projet Ariana

Rapport de recherche n° 5982 — Septembre 2006 — 14 pages

**Abstract:** We present a new model for extracting tree crowns from aerial images, containing a near infrared channel. We address the problem as an image segmentation problem solved by a Markov Random Field (MRF) modeling embedded in a Bayesian framework. The prior model consists of a couple field defined by a three state MRF, associated to vegetation, background and center of trees, and a template field representing the tree shape. A template field allows to associate an ellipse to model the crown of each detected tree. The model is optimized by a simulated annealing scheme based on a Metropolis dynamics. We show some results on real images of manmade plantations and natural forests.

**Key-words:** tree crown detection, Markov Random Field, couple field, template

This work was partially supported by EGIDE within the ECO-NET project 10203YK and the ARC "Mode de Vie" (INRIA). We thank the French Forest Inventory (IFN), the Center of Image Analysis of the Swedish University of Agricultural Science (SUAS) and the Royal Veterinary and Agricultural University of Denmark (RVAU) for providing the test images.

* E. Pechersky is within Lab. Dobrushin- IITP, Bol'shoi Karetnyi per., 19, 127994, GSP-4, Moscow, Russia, His work was partially supported by CNRS.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

# Extraction de houppiers par un champ de Markov à trois états

**Résumé :** Nous proposons une nouvelle approche pour extraire les houppiers d'arbre à partir d'images aeriennes contenant un canal proche infra-rouge. Nous résolvons le problème comme un problème de segmentation en utilisant une modélisation markovienne dans un contexte Bayesien. Le modèle a priori consiste en un champ couple défini par un champ de Markov à trois états, associés au fond, à la végétation et au centre des arbres, et par un champ de patrons représentant la forme des arbres. Le champ aléatoire de patrons permet d'associer un modèle elliptique à chaque houppier détecté. Le modèle est optimisé par un recuit simulé fondé sur une dynamique de Metropolis. Nous montrons des résultats sur des images réelles de plantations et de forêts naturelles.

**Mots-clés :** détection de houppiers, champs de Markov, champ couple, patron

# 1 Introduction

Tree counting and classification provide some parameters which are of paramount importance for managing the forest resources. An accurate characterization of the tree population is necessary to evaluate the evolution of the population or the effect of ecological disasters such as hurricanes or epidemics. Remotely sensed data play a leading role in this area as a cheap and fast way of investigation compared to an evaluation on the ground. Automatic analysis of remotely sensed images is another improvement as it saves expert time in making the evaluation. Early works in this area concerned essentially the detection of the tree position. As the image resolution increases, new algorithms can also provide an individual delineation of trees.

In this paper, we propose an approach to jointly estimate the tree position and delineate their crown. We consider three channel images as input, one of them being the near infrared channel. In this context, several approaches have been proposed for that purpose. A first class of approaches is based on a detection of seeds (representing candidates for representing a tree) which derives from the detection of local maximum above some threshold in the near infrared channel. The delineation of each individual tree can then be obtained by a contour approach based on a valley following procedure [1] or a region growing algorithm [2]. These approaches are based on the gray level information (trees induces light gray levels in near infrared images) and suppose that a shadow (dark gray levels) is visible around the trees. This last assumption can be partially ignored by applying post-treatment to split complex connected components. The second class of approaches consists of defining a template. A tree is detected if the image signal locally match the defined template [3, 4]. In this approach, the detection is not based on the pixel level but on a local mask. This approach is then more robust if the real shape of the tree crown well fit the template. To extend the generality of this approach, a set of different templates can be considered. The decision is then taken on a voting principle [5]. Extending the generality of the templates by adding some parameters may also lead to high computational time. Among this second class, a global model on the whole image can be defined by incorporating some information on the spatial structure of the population. Some regularity properties on the tree localization can be modeled by interactions between neighboring trees. A general model, based on marked point process has been recently proposed [6]. The solution is obtained by an optimization technique, based on a simulated annealing scheme [7]. In this paper, we combine the two previous classes. We propose a pixel based approach while incorporating some shape information as well as some interactions between different trees. We consider the Bayesian framework and define the problem as a segmentation problem. We define a Markov Random Field (MRF) as a prior. MRFs are an appropriate tool to address global constraints by defining local interactions [8, 9]. They have been widely used for image segmentation because of their ability to impose regularity constraints while preserving the details of shapes [10]. Herein, we consider some short and middle range interactions defined by a random template representing the shape of tree crowns and long range interactions defining the regularity of the tree localization. The data are taken into account at a pixel level by a Gaussian likelihood. Finally, the model optimization is performed by a simulated annealing scheme based on a Metropolis dynamics. One key point is that the interactions are defined by templates modeling trees on each pixel. The label field $L$ activates or not this template for each pixel, reflecting or not the presence of a tree at this location. This label field is not stationary in the sense that the templates may vary in space reflecting trees of different dimension. We therefore consider a second random field $T$ modeling the distribution of templates in the image. The couple $(L, T)$ is then a stationary couple Markov field as defined in [11].

The model is presented in section 2. Some results on real data are shown in section 3. Finally, conclusions are drawn in section 4.

## 2 Modeling tree population by a three states Markov Random Field

We extract the tree crown by segmenting the image using a Markov Random Field (MRF) approach embedded into a Bayesian framework. We first consider a set of 2D templates defined by two concentric ellipses, the first one representing the tree crown, the second one defining the background around the tree. Let $E$ denotes the template set:

$$\begin{aligned} & E = \{e_{(a,b,\theta)}, a \in [r_{min}, r_{max}], b \in [r_{min}, r_{max}], \theta \in [0, \pi/2[\}, \\ & \quad \text{where } e_{(a,b,\theta)} = e^{in}_{(a,b,\theta)} \cup e^{out}_{(a,b,\theta)} \\ & \text{with} \begin{cases} e^{in}_{(a,b,\theta)} = \{s = (i,j) \in S : \frac{(i\cos\theta + j\sin\theta)^2}{a^2} + \frac{(-i\sin\theta + j\cos\theta)^2}{b^2} \leq 1\} \\ e^{out}_{(a,b,\theta)} = \{s = (i,j) \in S, s \notin e^{in}_{(a,b,\theta)} \text{ and } \frac{(i\cos\theta + j\sin\theta)^2}{(a+2)^2} + \frac{(-i\sin\theta + j\cos\theta)^2}{(b+2)^2} \leq 1\} \end{cases} \end{aligned} \tag{1}$$

We define a random field $T$ on the lattice $S$, $t_s = s + e(s)$ where $e(s) \in E, \forall s \in S$. We now consider a realization of $T$, which means that on each pixel a template, which is an element of $E$ translated on $s$, is defined to model an underlying tree. Let $L$ be the label field where $l_s \in \{0, 1, 2\}, \forall s \in S$. The label 0 refers to background, the label 1 refers to vegetation (tree or not), and the label 2 refers to a tree center. The template $t_s$ at pixel $s$ is activated, which means that we consider that there is indeed a tree centered on $s$, if and only if $l_s = 2$. We embed the problem into a Bayesian framework and optimize the posterior $P(L, T|Y)$ where $Y$ represents the data:

$$P(L, T|Y) \propto P(Y|L, T)P(L, T) \propto P(Y|L, T)P(L|T)P(T) \tag{2}$$

The templates located at each site $s$ are supposed to be independent and uniformly distributed on $E$. We also assume that, conditionally on the label field, the data are independent from the template field. We then have to optimize the following expression:

$$P(L, T|Y) \propto P(Y|L)P(L|T) \tag{3}$$

The prior $P(L|T)$ is a Markov Random Field, defined as follows:

$$P(L|T) = \frac{1}{Z}\exp -\beta_t \sum_{s\in S} V(l_u, u \in t_s)\delta(l_s = 2) - \beta_r \sum_{(u,v)\in S\times S} \delta(l_u = 2, l_v = 2)\left|e^{in}(u) \cap e^{in}(v)\right|, \tag{4}$$

where $|A|$ denotes the cardinal of $A$. Note that the Markov property is satisfied as soon as the templates are bounded ($r_{max} < \infty$). The second term of equation (4) represents a repulsive term which penalizes overlapping between templates. The first term fits the label field on the templates and is defined as follows:

$$V(l_u, u \in t_s) = \sum_{u\in t_s^{in}} \delta(l_u = 0) - \delta(l_u = 1) + \sum_{u \in t_s^{out}} \delta(l_u = 1) - \delta(l_u = 0). \tag{5}$$

The template term thus favors vegetation labels in the template and background labels in the boundary. The likelihood is defined by modeling the background and the vegetation with Gaussian laws:

$$\begin{aligned} & P(Y|L,T) = \\ & \exp - \left[\sum_{s\in S} \left(\frac{(y_s-\mu_t) * C_t^{-1} *^t (y_s - \mu_t)}{2} + 0.5\log(2\pi Det(C_t))\right)(\delta(l_s = 1) + \delta(l_s = 2)) \right. \\ + & \left. \left(\frac{(y_s-\mu_b) * C_b^{-1} *^t (y_s - \mu_b)}{2} + 0.5\log(2\pi Det(C_b))\right)\delta(l_s = 0)\right] \end{aligned} \tag{6}$$

where $\mu_t, C_t$ (resp. $\mu_b, C_b$) are the mean and the covariance matrix of the vegetation pixels (or tree center) (resp. the background pixels). $Det(.)$ is the determinant.

To optimize the model we consider a MCMC dynamic embedded into a simulated annealing scheme. We consider the following Metropolis iterative algorithm:

a. Consider a random initial configuration $X^{(0)} = (L^{(0)}, T^{(0)})$, set $Temp = Temp(0)$ and $n = 0$

b. For each site $s \in S$, considered in a lexicographic order:

- select a new value $u$ according to a uniform law on $\{0, 1, 2\} \times E$,
- Define the current configuration $X = (x_t^{(n+1)}, t < s, x_s = x_s^{(n)}, x_t^{(n)}, t > s)$ and the new configuration $X' = (x_t^{(n+1)}, t < s, x_s = u, x_t^{(n)}, t > s)$ and compute the acceptance ratio $\alpha = \left(\frac{P(X'|Y)}{P(X|Y)}\right)^{\frac{1}{Temp(n)}}$
- set $x_s^{(n+1)} = u$ with probability $\min(1, \alpha)$ and $x_s^{(n+1)} = x_s^{(n)}$ with probability $1 - \min(1, \alpha)$

c. if the stopping criterion is not reached set $Temp(n+1) = f(n+1)$ and go to b

The function $f$ is a decreasing function. The convergence to the global maximum of the posterior distribution is theoretically obtained for $f(n) = \frac{Temp(0)}{\log(n+1)}$. In practice, to speed up the process, we use $f(n) = a^n Temp(0)$ with $a$ close to 1, typically $a = 0.98$.

## 3 Results

In this section, we show some results obtained on different kinds of tree population. First, let us consider a man made plantation, which consists of periodically spaced poplars. The data consist of a three channel aerial image with a 50 cm resolution. This first simple example validates the approach as the trees are correctly detected and localized (see figure 1). If we enlarge the area, we can see that, by defining a template which covers the object but also its neighborhood, and thus by favoring background pixels in the neighborhood, we avoid false alarms in the fields, where the radiometry is close to that of the trees (see figure 2). In this second example, a few false alarms occur on the right handside of the plantation. This is due to both the presence of a field with a radiometry similar to the trees and to shadows of the trees at the border of the plantation. Note that we do not introduce any information concerning the tree alignment in the model. Therefore, the detection remains robust in more difficult cases as shown on figure 3. On figure 3 bottom right, we show the intersection of the activated templates and the pixels labeled as vegetation. These connected components represent the actual detected tree crown shapes. They can be used for example to compute an estimation of the total crown surface, which can be linked to the total wood volume when the tree species is known.

The next example concerns a Swedish forest, in which four different species can be distinguished (pine, spruce, birch, aspen). Despite the different typical sizes and some radiometry variation, the detection is still satisfactory, mainly because the mode representing the background is well defined (see figure 4).

We now address the limits of the approach. Since we have proposed a template based approach, the performances depend on the geometrical model embedded in the template definition. We show the limit of the approach on two different examples. The first one represents a very dense area (see figure 5). In this area, the contour of each crown is hardly visible, even for an expert. The detection shows that only part of the scene is described. The overlay of the detected ellipses shows that there is not a clear correspondence between the data and the detected trees. Our object model is not adapted to very dense areas for which only a global description can be derived from the information observed in the data.

Finally, we show a result on an image taken far from the nadir point (see figure 6). The perspective effect spoils the performance of our approach. Only the main trees have been correctly detected. For this kind of images, a fully template based approach such as in [4] is more appropriate. Such an approach, considering a wider set of possible templates, is difficult to embed in a stochastic framework, due to computational considerations.

# 4 Conclusion

We have proposed a couple Markov random field to extract and delineate tree crowns from aerial images. This model consists of a Markov label field containing middle range interactions defined by the realization of an underlying template field. Results have been shown on several datasets, proving the efficiency of the approach. This model can be seen as a compromise between an object approach (defined by the template field) and a pixel approach (defined by the label field). Therefore, the model reaches its limits when the templates do not fit the data, for instance when the population is too dense. The next step will consist in comparing and evaluating different approaches for tree extraction with respect to detections performed by an expert. We are currently working on a protocol to conduct this evaluation. A second aspect concerns tree classification. When the tree crown is extracted, we can expect to classify the different species using some features based on radiometry, texture or/and shape. This can provide information for studying natural resources and give for example some estimation of the wood volume contained in a given population.

# References


[1] F.A. Gougeon, "Automatic Individual Tree Crown Delineation using a Valley-following Algorithm and Rule-based System," in *Proc. of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, D.A. Hill and D.G. Leckie, Eds., Victoria, British Columbia, Canada, February 1998, pp. 11–23.

[2] M. Erikson, *Segmentation and Classification of Individual Tree Crowns*, Ph.D. thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden, 2004.

[3] R. Pollock, "Individual Tree Recognition based on Synthetic Tree Crowns Image Model," in *Proc. of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, D.A. Hill and D.G. Leckie, Eds., Victoria, British Columbia, Canada, February 1998, pp. 25–34.

[4] M. Larsen, "Crown modelling to find tree top positions in aerial photographs," in *Proc. of the Third Internationl Airborne Remote Sensing Conference and Exhibition*, 1997, vol. 2, pp. 428–435.

[5] M. Larsen, "Individual Tree Top Position Estimation by Template Voting," in *Proc. of the Fourth International Airborne Remote Sensing Conference and Exhibition / $21^{st}$ Canadian Symposium on Remote Sensing*, Ottawa, Ontario, June 1999, vol. 2, pp. 83–90.

[6] G. Perrin, X. Descombes, and J. Zerubia, "2D and 3D vegetation resource parameters assessment using marked point processes," in *Proc. International Conference on Pattern Recognition (ICPR)*, Hong-Kong, August 2006.

[7] G. Perrin, X. Descombes, and J. Zerubia, "Adaptive simulated annealing for energy minimization problem in a marked point process application," in *Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*, St Augustine, Florida, USA, November 2005.

[8] J. Besag, "Spatial interaction and statistical analysis of lattice systems," *Journal of the Royal Statistical Society Series B*, vol. 36, pp. 721–741, 1974.

[9] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images.," *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[10] X. Descombes, R. Morris, J. Zerubia, and M. Berthod, "Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood," *IEEE Trans. Image Processing*, vol. 8, no. 7, pp. 954–963, July 1999.

[11] W. Pieczynski and A.N. Tebbache, "Pairwise Markov random fields and segmentation of textured images," *Machine Graphics and Vision*, vol. 9, no. 3, pp. 705–718, 2000.

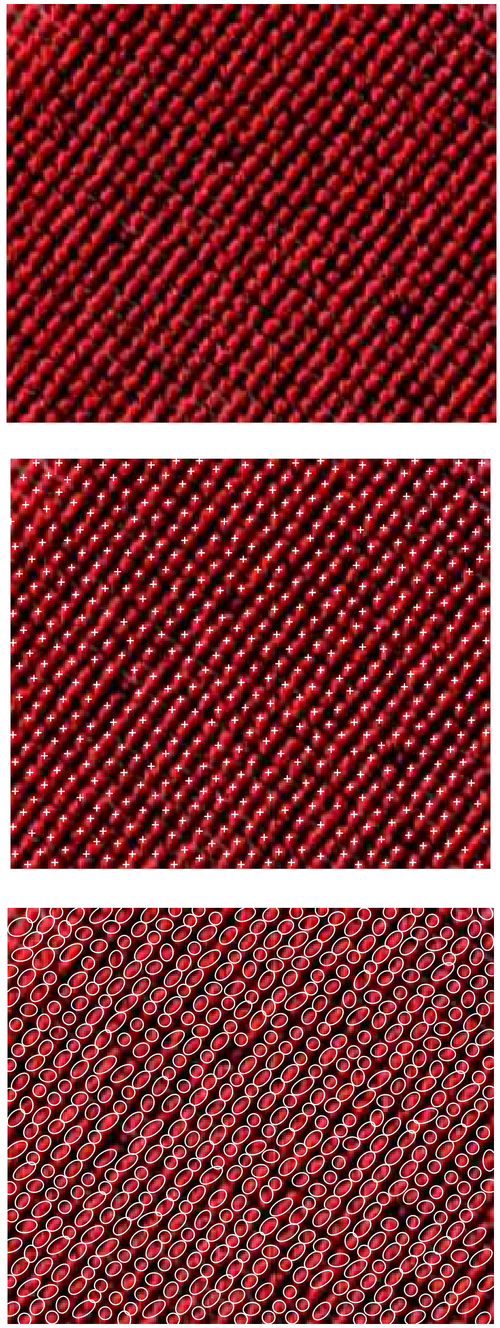

Figure 1: Result on a poplar plantation (top: initial image © IFN, middle: detected trees, bottom: tree templates)

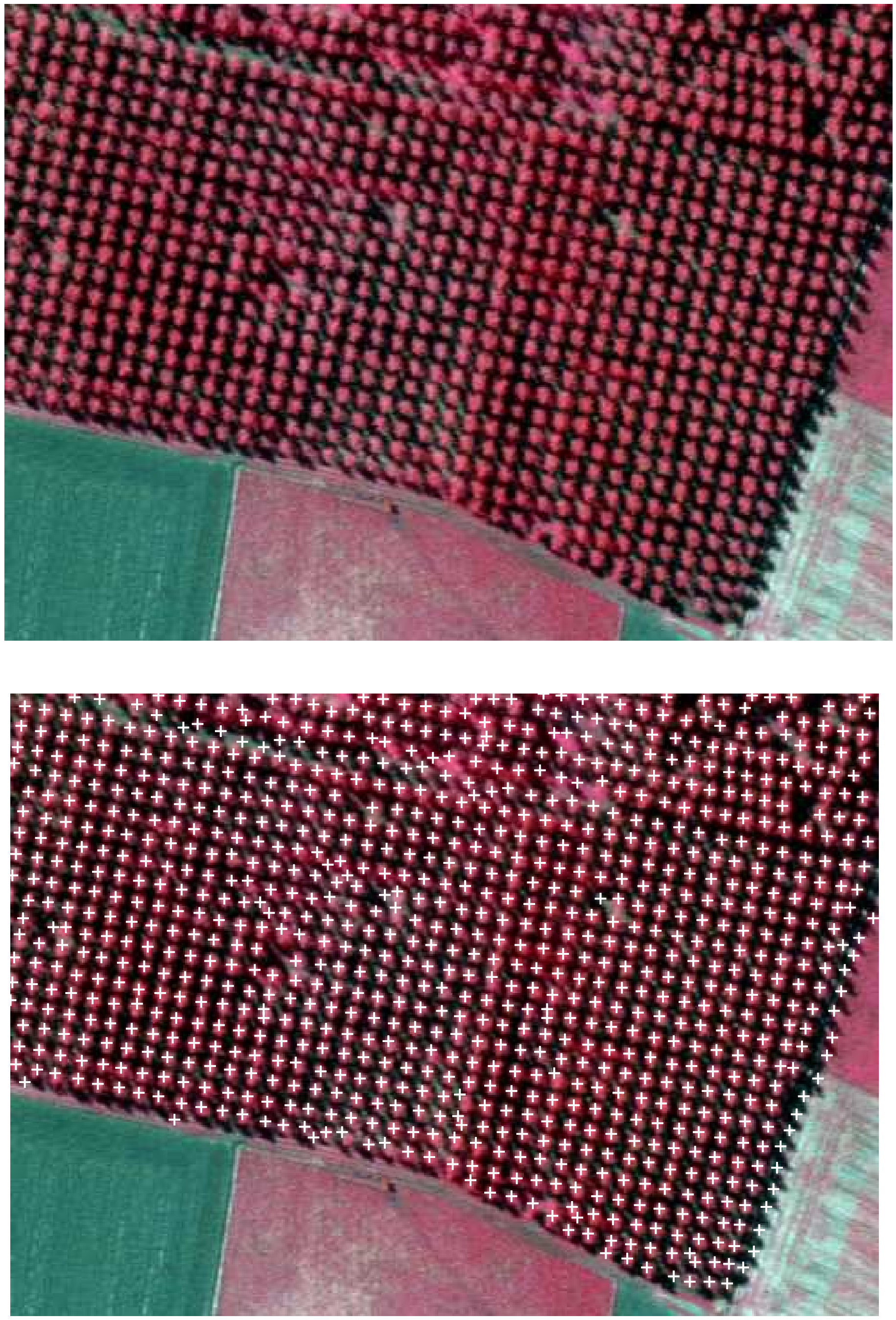

Figure 2: Result on a poplar plantation (top: initial image © IFN, bottom: detected trees)

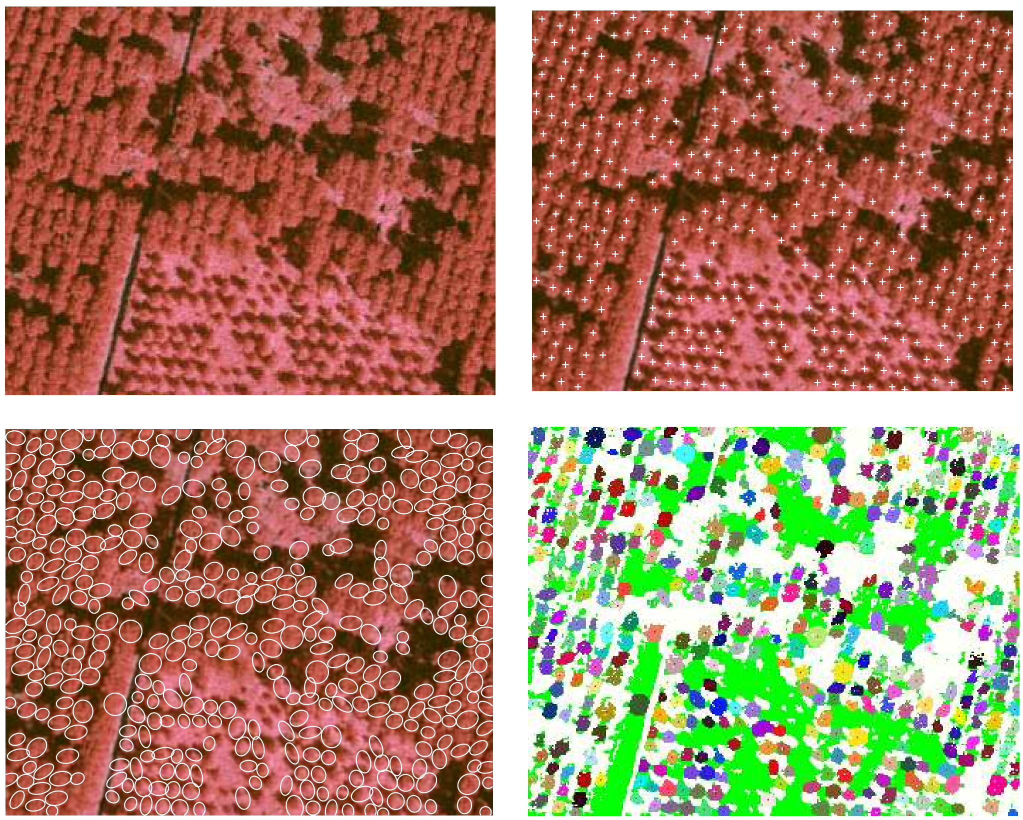

Figure 3: Result on a poplar plantation (top left: initial image © IFN, top right: detected trees, bottom left: tree templates, bottom right: detected crown shape.)

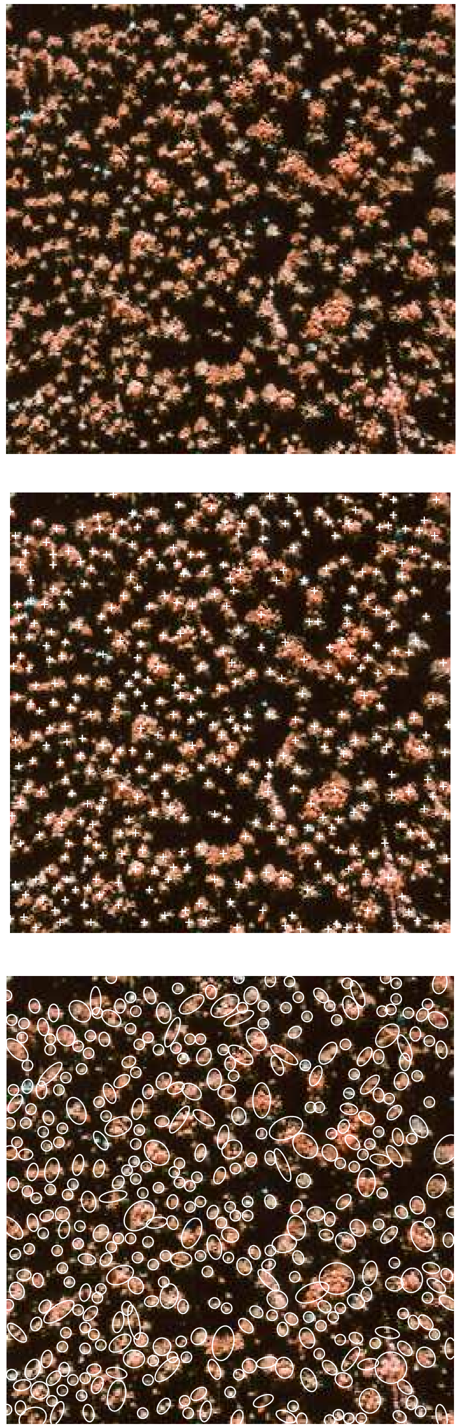

Figure 4: Result on a natural forest (top: initial image © SUAS, middle: detected trees, bottom: tree templates)

Figure 5: Result on a dense oak forest (top: initial image © IFN , middle: detected trees, bottom: tree templates)

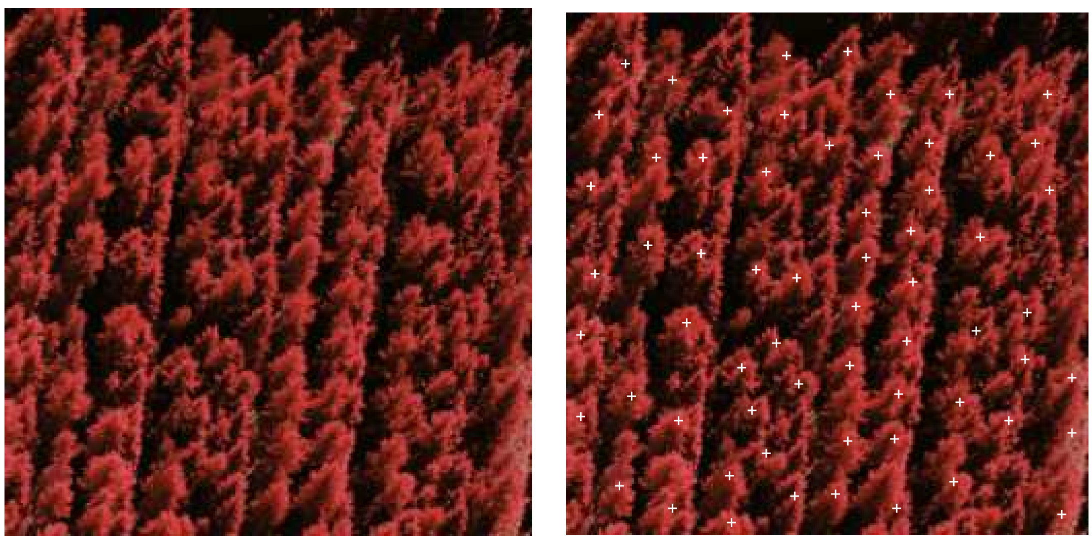

Figure 6: Result on a image far from the nadir point (top: initial image © RVAU, bottom: detected trees)

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399